Final Exam for CS 5354/CS 4365, Fall 2022

1. *General:*
   a. explain what is the main objective of this class;
   b. what did your group do as a project and how does it fit into the general objective of the class.
2. *Invariant functions:*
   a. define shift-shift invariance;
   b. prove that every shift-shift invariant function is linear;
   c. describe the general forms of shift-scale, scale-shift, and scale-scale invariant functions (no proof is needed);
   d. describe the function $y = \ln(1 + \exp(x))$ as a composition of invariant functions.
3. *Invariant families of functions:*
   a. why do we need to consider families of functions? what is the general form of families that we consider?
   b. describe all functions from a shift-invariant family (no proof is needed);
   c. describe all functions from a scale-invariant family (no proof is needed);
   d. describe all functions from a family which is both shift-invariant and scale-invariant (no proof is needed);
   e. use invariance to explain a function $\exp(x) + 2x \cdot \exp(-x)$.
4. *Optimality:*
   a. use calculus to find the point x at which the function $y = x^2 - 2x + 1$ attains its minimum; what is the value of y at this point?
   b-c. prove that if an optimality criterion is final and scale-invariant, then the optimal alternative is also scale-invariant; explain what *final* means;

d. describe all functions from a family that is optimal with respect to a final scale-invariant optimality criterion.

5. *Transformation groups:*

   a. explain what is a transformation group;

   b. explain why the class of all natural transformations should be a finite-parametric transformation group that contains all linear transformations;

   c. prove that the class of all shifts is a transformation group;

   d. describe all functions from a finite-parametric transformation group that contains all linear transformations;

   e. use invariance to explain a function $y = \dfrac{x}{a + x}$.

6. Neural networks:

   a. what is the general formula describing a neuron?

   b. how can you explain the activation function $s(x) = 1/(1 + \exp(-x))$ used in traditional (shallow) neural networks?

   c. use scale-invariance to explain the ReLU activation function $s(x) = \max(0, x)$;

   d. why do we need pooling in neural networks?

   e. which pooling operations are shift-and scale-invariant? (no proof needed)